Pivotal Tuning for Latent-based Editing of Real Images

Abstract

Recently, numerous facial editing techniques have been proposed that leverage the generative power of a pretrained StyleGAN. To successfully edit an image this way, one must first project (or invert) the image into the pretrained generator's domain. As it turns out, StyleGAN's latent space induces an inherent tradeoff between distortion and editability, i.e., between maintaining the original appearance and convincingly altering its attributes. Hence, it remains challenging to apply ID-preserving edits to real facial images. In this article, we present an approach to bridge this gap. The idea is Pivotal Tuning—a brief training process that preserves editing quality, while surgically changing the portrayed identity and appearance. In Pivotal Tuning Inversion, an initial inverted latent code serves as a pivot, around which the generator is fine-tuned. At the same time, a regularization term keeps nearby identities intact, to locally contain the effect. We further show that Pivotal Tuning also applies to accommodating for a multitude of faces, while introducing negligible distortion on the rest of the domain. We validate our technique through inversion and editing metrics and show preferable scores to state-of-the-art methods. Last, we present successful editing for harder cases, including elaborate make-up or headwear.